UNCONSTRAINED FACE IDENTIFICATION WITH MULTI-SCALE BLOCK-BASED CORRELATION

Jack Gaston, Ji Ming, Danny Crookes

Institute of Electronics, Communications and Information Technology
Queen's University Belfast, Belfast BT3 9DT, UK

ABSTRACT

Many approaches to unconstrained face identification exploit small patches which are unaffected by distortions outside of their locality. However, small patches have limited discriminative ability, making accurate patch matching difficult. We propose a novel *block*-based approach to exploit the greater discriminative information in larger areas, while maintaining robustness to local variations. A testing *block* contains several neighbouring testing patches. We identify all the matching training patches in a block *jointly*, using normalized cross correlation (NCC), as a means of reducing the uncertainty of each matching patch with the addition of the neighbouring patch information. We further propose a multi-scale extension in which we carry out block-based matching at several block sizes, where a larger block contains more neighbouring testing patches, to combine complementary information across scales for further robustness. For evaluation, we use two unconstrained datasets, cropped Labelled Faces in the Wild (LFWCrop) and Unconstrained Facial Images (UFI). Our new approach is able to significantly improve identification accuracy over existing patch-based methods, in the presence of uncontrolled pose, expression and lighting variations.

Index Terms— Face identification, multi-patch blocks, block-based correlation, multi-scale approach, unconstrained variations.

1. INTRODUCTION

Face identification performance degrades in real-world scenarios. Variations caused by pose and illumination changes can be greater than those caused by a subject's identity [1]. Because these variations affect different regions differently, many approaches treat images as ensembles of *small* patches [2–21], assuming that each patch is unaffected by variations outside its locality. Recognition can be performed by finding a matching training patch for each testing patch. In many current methods, this is done *independently* for each patch. For example, sparse representation classification (SRC) forms an estimate of each testing patch, as a linear combination of the training patches, independently of the other testing patches; the independent patch classifications can then be combined in a voting scheme, for final classification [2]. More recently, a matching scheme using normalized cross correlation (NCC), or Pearson's correlation coefficient, was applied to unconstrained face identification [3], which identifies the matching patches independently based on maximum patch-sized correlation. Variations can also be handled by including or synthesizing diverse conditions in training data. This has been used for patch-based and holistic SRC [2, 22, 23], as well as collaborative representation classification (CRC) [24].

Patch-based approaches can also be used to deal with variable lighting conditions. By assuming piecewise constant lighting across the patches [4, 25, 26], illumination normalization filters such as self quotient images (SQIs) [5] and gradientfaces [6] can be applied in small patches, each with a near even lighting condition, to remove low frequency illumination variation while maintaining fine details. Local binary patterns (LBPs) [7], and its variants [8–14] , describe small patches, assuming constant or low-frequency lighting in each patch, by encoding high frequency variations which represent facial structure. The NCC metric can cancel the constant lighting difference between small patches for comparison(e.g., [3, 26]).

Small patches can be limited in their discriminative ability, making their *independent* matching prone to errors. However, larger matching areas may be unidentifiable due to local appearance changes across the region. Multi-scale approaches combine the complementary information at several patch-sizes to mitigate this problem. Multi-scale LBPs (MLBP) [15, 16] concatenate LBPs at several patch-sizes to improve robustness, the hierarchical MLBP (HMLBP) reduces feature dimensionality while maintaining discriminative information by selecting a single appropriate scale for each image location [17]. The pattern of oriented edge magnitudes (POEM) [18] is a multi-scale descriptor where local appearance is captured by accumulating gradient magnitudes in small patches, and the larger structure is captured by extracting LBPs from larger blocks in the derived gradient images. Multi-scale extensions to subspace-based approaches, including SRC (MSRC) [19] and CRC (MPCRC) [20, 21], have also improved robustness.

In this paper, we extend the previous work towards a more robust approach to unconstrained facial image identification. Our extension includes two novel aspects. First, we propose a *block*-based approach, in place of the conventional *patch*-based approach, to improve discrimination for face identification. A testing block contains a number of neighboring patches which are matched *jointly* to reduce the uncertainty for each individual matching training patch. By optimising both the matching training patches and their illumination difference with the corresponding testing patches, we exploit the discriminative information in larger regions while maintaining robustness to local variations. Second, we propose a multi-scale extension to our block-based matching, where a block at a higher scale contains more neighbouring testing patches, to further enlarge the matching areas and combine their complementary information to improve discrimination and robustness. We have compared our new approach with existing approaches on two difficult unconstrained facial databases and achieved improved performance.

2. MULTI-SCALE BLOCK-BASED CORRELATION

2.1. Block Correlation

Our idea for improving facial identification accuracy can be illustrated with Fig. 1. Assume that we compare a testing image against a training image from person s, and that the testing image is divided

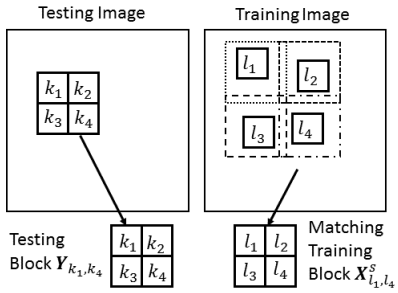

Fig. 1: The selection of 4 matching training patches for a testing block of size 2 patches by 2 patches. Matches are found *jointly*, within search windows, indicated by dotted lines, to allow for local displacement and enforce a semantic constraint on matching patches.

into small patches located by the patch indexes $k_1, k_2, ...$ Conventional patch-based methods tend to process these patches independently. In our new method, we aim to match a number of neighboring patches *jointly*. For example, in Fig. 1 we show the match of a *block* of 2 patches by 2 patches, k_1, k_2, k_3 and k_4. By considering a larger block area we will normally gain greater discriminative ability, and hence reduce uncertainty of the individual matching training patches, compared to the patch-based method. However, because of the potential pose/expression variations between unconstrained images, large, perfectly matching blocks may not exist. We assume that these local variations can be modeled by allowing for certain local displacement of the matching training patches, as an approximate model for the lack of pose/expression/alignment constraints on the data, but maintaining a similar geometric relationship, as a semantic constraint on valid facial structure, as Fig. 1 illustrates.

Let $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_K)$ represent a testing image divided into K small, partially overlapped patches, where $\mathbf{y}_k$ is the kth patch, with origin (top-left corner coordinates) (i_k, j_k). For simplicity, we assume that all patches have the same size: $M \times N$ pixels. We apply the same representation to the training images such that the training image of person s, X^s, has an expression $\mathbf{X}^s = (\mathbf{x}^s_1, \mathbf{x}^s_2, ..., \mathbf{x}^s_L)$, where $\mathbf{x}^s_l$ is the lth $M \times N$-sized patch. One person can have more than one training image. We use the NCC as a comparison metric and identify matching *blocks* to improve identification accuracy. One reduced form of our approach, described in [3], identified the individual matching *patches* independently. The NCC of two patches, $\mathbf{y}_k$ and $\mathbf{x}^s_l$, is measured with the expression:

$$R(\mathbf{y}_k, \mathbf{x}^s_l) = \frac{\sum_{m,n}[y_k(m,n) - \mu_k][x^s_l(m,n) - \mu^s_l]}{\sigma_k \sigma^s_l} \quad (1)$$

where $y_k(m,n)$ denotes a pixel at (m,n) inside patch $\mathbf{y}_k$, the sum is over all pixels in the patch, μ_k is the mean pixel value of patch $\mathbf{y}_k$, and σ_k is the zero-mean Euclidean norm of patch $\mathbf{y}_k$, i.e., $\sigma^2_k = \sum_{m,n}[y_k(m,n) - \mu_k]^2$. The same definitions apply to the training patch $\mathbf{x}^s_l$ with mean pixel value μ^s_l and zero-mean Euclidean norm σ^s_l. In [3], given a testing image, the matching training patches were identified independently, by maximizing $R(\mathbf{y}_k, \mathbf{x}^s_l)$ for each testing patch $\mathbf{y}_k$ over the possible matching training patches $\mathbf{x}^s_l$. In this paper, we extend this approach to identifying larger matching blocks, as illustrated in Fig. 1, to improve discrimination. Let $\mathbf{Y}_{k_1,k_Q} = (\mathbf{y}_{k_1}, \mathbf{y}_{k_2}, ..., \mathbf{y}_{k_Q})$ represent a testing block with Q patches (block $\mathbf{Y}_{1,K}$ contains all patches in the image and hence is equivalent to the whole image $\mathbf{Y}$), and let $\mathbf{X}^s_{l_1,l_Q} = (g_{l_1}\mathbf{x}^s_{l_1}, g_{l_2}\mathbf{x}^s_{l_2}, ..., g_{l_Q}\mathbf{x}^s_{l_Q})$ represent a possible matching training block, where the individual matching training patch locations l_q $(q = 1, 2, ..., Q)$ may be a nonlinear function of the corresponding testing patch locations k_q to account for the potential nonlinear local displacement between the training and testing images; g_{l_q} is the gain for each local training patch, to compensate for lighting differences. We can assume that g_{l_q} is a constant within each small patch, but can be variable across the block. This is identical to a piece-wise constant lighting model [4, 25, 26]. The NCC for the two blocks can be written as:

$$R(\mathbf{Y}_{k_1,k_Q}, \mathbf{X}^s_{l_1,l_Q}) = R(\mathbf{y}_{k_1}\mathbf{y}_{k_2}...\mathbf{y}_{k_Q}, g_{l_1}\mathbf{x}^s_{l_1}g_{l_2}\mathbf{x}^s_{l_2}...g_{l_Q}\mathbf{x}^s_{l_Q})$$
$$= \frac{\sum_{q=1}^{Q} g_{l_q}\sum_{m,n} y_{k_q}(m,n)x^s_{l_q}(m,n) - T\mu_{k_1,k_Q}\mu^s_{l_1,l_Q}}{\sigma_{k_1,k_Q}\sigma^s_{l_1,l_Q}} \quad (2)$$

where $T = QMN$ is short-hand notation for the total number of pixels in each block, $\mu^s_{l_1,l_Q}$ and $\sigma^s_{l_1,l_Q}$ are the global mean pixel value and zero-mean Euclidean norm of training block $\mathbf{X}^s_{l_1,l_Q}$, i.e.,

$$\mu^s_{l_1,l_Q} = \frac{1}{T}\sum_{k=q}^{Q} g_{l_k}\sum_{m,n} x^s_{l_k}(m,n) \quad (3)$$

$$[\sigma^s_{l_1,l_Q}]^2 = \sum_{k=1}^{Q} g^2_{l_k}\sum_{m,n}[x^s_{l_k}(m,n)]^2 - T[\mu^s_{l_1,l_Q}]^2 \quad (4)$$

The above (3) and (4) apply to μ_{k_1,k_Q} and σ_{k_1,k_Q}, the global mean and zero-mean Euclidean norm of testing block $\mathbf{Y}_{k_1,k_Q}$ (without the gain terms). Unlike previous approaches (e.g., the single patch based NCC [3], and SRC [19]) in which the individual patches are treated independently, in the proposed block-based NCC (2) there is no assumption of independence between the patches in a block. We will use (2) to identify the matching patches *jointly*.

2.2. Optimal Block Matching

Given a testing block $\mathbf{Y}_{k_1,k_Q}$, we seek the best concatenation of training patches to estimate a matching training block. To constrain the estimate to valid facial blocks, we formulate the estimation as a constrained maximization problem. We give a general expression:

$$R^{\mathcal{L}}(\mathbf{Y}_{k_1,k_Q}, \hat{\mathbf{X}}^s_{l_1,l_Q})$$
$$= \max_{\mathbf{X}^s_{l_1,l_Q}} R(\mathbf{Y}_{k_1,k_Q}, \mathbf{X}^s_{l_1,l_Q})\mathcal{L}(\mathbf{X}^s_{l_1,l_Q}|\mathbf{Y}_{k_1,k_Q})^\alpha \quad (5)$$

where $\hat{\mathbf{X}}^s_{l_1,l_Q}$ represents the optimal estimate of the matching training block, found from the training data for person s. The optimal estimate has maximum block-based NCC subject to a constraint defining valid facial blocks, represented by the multiplying likelihood function $\mathcal{L}(\mathbf{X}^s_{l_1,l_Q}|\mathbf{Y}_{k_1,k_Q})^\alpha$, that $\mathbf{X}^s_{l_1,l_Q}$ is likely to be a valid matching block for the given testing block $\mathbf{Y}_{k_1,k_Q}$, where α denotes the weight for the constraint; $R^{\mathcal{L}}(\mathbf{Y}_{k_1,k_Q}, \mathbf{X}^s_{l_1,l_Q})$ represents the constrained NCC and the value associated with the optimal estimate is $R^{\mathcal{L}}(\mathbf{Y}_{k_1,k_Q}, \hat{\mathbf{X}}^s_{l_1,l_Q})$.

As an example, in our experiments, we use a simple semantic constraint on facial structure, by enforcing that the selected training patches have a similar geometric relationship to their corresponding testing patches. Specifically, we assume that the location of the matching training patch l_q corresponding to testing patch k_q can be expressed as a function of k_q as follows:

$$l_q = k_q + h + \delta_q \quad (6)$$

where h represents a common displacement for all patches in the block, which may be caused by some major misalignment which

affects all the patches equally, and δ_q represents a local displacement for a specific patch, which may be caused, for example, by random local pose or expression changes. Based on (6), we use a simple expression for the likelihood function:

$$\mathcal{L}(\mathbf{X}^s_{l_1,l_Q}|\mathbf{Y}_{k_1,k_Q}) = \prod_{q=1}^{Q} I_{\Delta_q}[l_q - (k_q + h)] \quad (7)$$

where $I_{\Delta_q}[l_q - (k_q + h)]$ is an indicator function of the form

$$I_{\Delta_q}[l_q - (k_q + h)] = \begin{cases} 1 & \text{if} \quad ||l_q - (k_q + h)|| \leq \Delta_q \\ 0 & \text{if} \quad ||l_q - (k_q + h)|| > \Delta_q \end{cases} \quad (8)$$

where Δ_q is a threshold on the small local displacement δ_q. Some displacement is allowed to account for local variations in pose, expression and alignment, but an upper limit is used as a semantic constraint, requiring that the matching training patches maintain a similar geometric relationship to the corresponding testing patches.

We use a computationally efficient iterative algorithm to solve the above constrained maximization problem (5). Given a testing block $\mathbf{Y}_{k_1,k_Q} = (\mathbf{y}_{k_1}, \mathbf{y}_{k_2}, ..., \mathbf{y}_{k_Q})$, we find an initial estimate $\hat{\mathbf{X}}^s_{l_1,l_Q} = (\hat{g}_{l_1}\hat{\mathbf{x}}^s_{l_1}, \hat{g}_{l_2}\hat{\mathbf{x}}^s_{l_2}, ..., \hat{g}_{l_Q}\hat{\mathbf{x}}^s_{l_Q})$ by separately estimating each matching training patch $\hat{\mathbf{x}}^s_{l_q}$ by maximizing the patch-based NCC $R(\mathbf{y}_{k_q}, \mathbf{x}^s_{l_q})$ with a unit gain $\hat{g}_{l_q}$. Then we update this initial estimate by alternately re-estimating each matching training patch, with gain, to maximize the block-based constrained NCC $R^{\mathcal{L}}(\mathbf{Y}_{k_1,k_Q}, \mathbf{X}^s_{l_1,l_Q})$; in re-estimating a specific training patch, the other training patches are fixed to their latest estimates. This alternate re-estimation process iterates until convergence is achieved. This algorithm manages to estimate the optimal training patches one patch at a time, subject to the constraints of all the other patches in the block, and hence can be calculated efficiently. It is shown that this algorithm converges in terms of generating a block estimate that always increases the block-sized, constrained NCC with each iteration. More details are given below.

2.3. Multi-Scale Block-Based Face Identification

Our approach can be extended to include scores for variable block-sizes, where block-size is defined by the number of patches in a block, to combine the complementary information across scales. Our block-based approach maintains invariance to local variations, even at large block-sizes, by optimizing the constituent patches. For example, at the testing block location k_1, we identify a number of matching training blocks for the testing blocks $\mathbf{Y}_{k_1,k_Q}$ with variable sizes Q, using (5), and then calculate the overall matching score between these blocks using the following expression

$$Score_s(k_1) = \sum_{Q} R^{\mathcal{L}}(\mathbf{Y}_{k_1,k_Q}, \hat{\mathbf{X}}^s_{l_1,l_Q}) w_Q \quad (9)$$

where w_Q denotes the weight for block size Q, reflecting its importance in the combination. The overall score for person s is formed by summing $Score_s(k_1)$ over all the testing block locations k_1. In our experiments, we have found that this multi-scale approach produces superior results to any individual block size, indicating its ability to capture complementary information from the variable-sized blocks.

3. EXPERIMENTAL EVALUATION

3.1. Experimental Settings

We assessed the efficacy of our approach for face identification in unconstrained conditions using the cropped Labelled Faces in the Wild (LFWCrop) [27] and Unconstrained Facial Images (UFI) [28] datasets. The LFWCrop dataset [27], contains pre-cropped images resized to 64×64 pixels, but still exhibiting real-life conditions, including misalignment. On LFWCrop we used a subset of 86 subjects with 11-20 images each, previously used in [22]. As in [22], we used 7 images per subject for training, and the rest for testing. The UFI set contains 605 subjects with a single testing image each, and an average of 7.1 training images. We used the cropped version for our experiments, and resized the images to 64×64 pixels. Example LFWCrop and UFI images are shown in Fig. 2. Throughout our NCC-based experiments, we used 8×8 patches with an overlap of 4 pixels. In the cropped images the facial features still exhibit various degrees of local variation, but occupy similar image regions (e.g. Fig. 2). Therefore, we assume that the common patch displacement $h = 0$ in (6), and the local patch displacement δ_q for each matching training patch is subject to a 30×30 search window surrounding the corresponding testing patch, i.e., $\Delta_q = 15$ in (8). In our approach matching patches may be selected from any of the training images for a subject, and the gain value for each matching training patch, g_{l_q}, is estimated by golden section search within the range $[0.1, 3]$. Unless otherwise indicated, we weighted scores of blocks from 1×1 up to 3×3 patches on LFWCrop, and from 1×1 up to 4×4 patches on UFI, with the weights tuned for each dataset, where smaller scales generally receive higher weights.

Fig. 2: Top: A testing image and 7 training images from LFWCrop. Bottom: A testing image and 7 training images from UFI.

3.2. Experimental Results on LFWCrop and UFI

On the LFWCrop dataset we compare our new Multi-scale Block-based NCC (M-BNCC) approach to a number of recent approaches to unconstrained face identification with the same training/testing subsets, particularly local patch-based and multi-scale techniques. These include a recent extension to the SRC framework [22], the HMLBP approach [17], the MPCRC [20] approach, and finally, the original patch-based NCC matching scheme [3]. We obtained the results for these approaches by running the code provided by the authors, except for the patch-based NCC which may be viewed as a special case of our M-BNCC approach. Table 1 shows the identification accuracy for each approach, indicating that our M-BNCC approach comfortably outperforms the existing approaches. As reported in [22], the SRC approach achieved identification accuracy of 34.36% on the same sets of LFW images we use, but manually cropped and resized. The performance drop may be caused by the uncontrolled misalignment in LFWCrop images. As reported in [20], MPCRC was able to achieve accuracy of 49% on LFW subsets with pre-alignment, unlike the mis-aligned LFWCrop images.

Table 1: Identification rate (%) on LFWCrop, comparing our new M-BNCC approach with others.

SRC [22]	MPCRC [20]	HMLBP [17]	NCC [3]	M-BNCC
28.37	15.04	36.28	48.37	**65.27**

Table 2: Identification rate (%) on UFI, comparing our new M-BNCC approach with others.

LBP [7]	LDP [13]	FS-LBP [14]	POEM [18]	M-BNCC
55.04	50.25	63.31	67.11	**74.55**

As reported in [3], on a subset of 50 LFW subjects, each with 5 training images and 3 testing images, and after applying the pose and illumination normalization process from [29], the original patch-based NCC scheme and the HMLBP approach achieved identification rates of 54% and 47% respectively. However, on a larger dataset than this, with no normalization process, we are able to produce higher identification rates using our M-BNCC approach, as shown in Table 1. Some other techniques have presented superior identification accuracy on various subsets of the LFW database, but make use of additional external data to improve performance (e.g., in [30] discriminative similarity measures were learned, exploiting negative background examples not seen in the database, to boost SVM classification and in [31] a probabilistic learning approach used weakly labelled data from internet searches to increase the amount of training data) or use many more training images per subject (e.g., [12]).

On the more recent UFI dataset we compare with the baseline approaches presented in [28]. The baseline UFI results include the traditional LBP approach [7], and the more recent Local Derivative Patterns (LDPs), which extracts LBPs in high order derivative space [13], Face-Specific LBP (FS-LBP) [14], where facial feature points are identified to extract LBPs from, rather than extracting LBPs in a regular grid, and the multi-resolution POEM descriptor [18] which achieved the best baseline performance. These results are cited from [28] and presented in Table 2, along with our M-BNCC results. To the best of our knowledge, our approach has superior performance to the best published results on this unconstrained dataset.

3.3. Further M-BNCC Algorithmic Evaluation

For the new M-BNCC approach, we evaluate the importance of the joint estimation of the matching patches in each block, and the semantic constraint to confine the matching patches' locations, towards improving the identification accuracy. Table 3 shows the identification accuracy on the LFWCrop dataset, at several block-sizes, with the joint estimation or without joint estimation (i.e., the matching patches are searched independently of one another), and with the semantic constraint or without this constraint (i.e., the matching patches can come from anywhere in the training images). We can see that our joint estimation significantly improves the discriminative ability, in terms of identification accuracy. Our semantic constraint on the location of selected matching training patches also improves identification accuracy. In Table 4 we present the identification accuracy with and without our estimated gain factors, g_{l_q}, showing that they attenuate the effect of local illumination variation on the selection of a matching block, increasing identification accuracy. At larger block-sizes, where non-uniform lighting is more likely, the benefit of optimised gain factors is more pronounced.

Table 3: Identification rate (%) on LFWCrop based on variable-sized blocks, e.g., 2 patches $\times$ 2 patches, showing the importance of joint estimation of the matching patches and the semantic constraint.

Estimation / Constraint	2×2	3×3	4×4
Independent / Unconstrained	19.22	20.00	21.55
Joint / Unconstrained	53.18	50.39	47.75
Independent / Constrained	24.96	25.27	25.89
Joint / Constrained	63.87	60.62	59.22

Table 4: Identification rate (%) on LFWCrop at 3 block-sizes, with optimised gain factors, g_{l_q}, and without gain optimisation ($g_{l_q} = 1$)

Gain value	2×2	3×3	4×4
$g_{l_q} = 1$	56.74	52.40	47.13
Optimised g_{l_q}	63.87	60.62	59.22

Table 5 shows the benefit of our multi-scale combination, across 4 block-sizes. We can see that each scale included in the weighted combination incrementally improves the identification accuracy by providing complementary information. We note that different scales classify different images correctly, and that a multi-scale combination can correctly classify face images that none of the constituent scales could. Finally, Fig. 3 summarizes the convergence of the iterative algorithm, showing the average number of iterations, and the initial and final constrained block-based NCC scores, on LFWCrop, at 4 block-sizes. Our search for the optimal matching block increases the NCC at each iteration. We stop the iteration when none of the block's constituent patches are updated between iterations.

Table 5: Identification rate (%) on UFI by the proposed M-BNCC, showing the combination of 4 block-sizes, from 1×1 up to 4×4, and the contribution of each additional scale to the final accuracy.

1×1	$+2 \times 2$	$+3 \times 3$	$+4 \times 4$
72.56	73.72	74.05	74.55

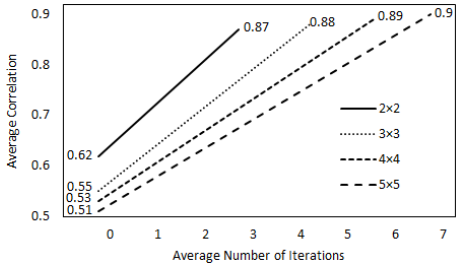

Fig. 3: Increase in constrained block correlation with iteration, at 4 block-sizes.

4. CONCLUSION

This paper studied the problem of face identification with unconstrained pose, expression and lighting variations. We extended the previous patch-based NCC matching scheme to block-based matching. A block contains several neighbouring patches. We identify all the matching patches in a block jointly to reduce the uncertainty in identifying each individual matching patch. By estimating the lighting difference between small matching patches, we attenuate the effect of illumination variation across the block. Finally, by combining our block-based classification scores at several block-sizes, we combine the complementary information across scales for further robustness. On two unconstrained datasets we have shown the benefit of our contributions, and improved identification accuracy significantly over existing patch-based methods. In future, we would like to develop an algorithm for learning weights to combine scales more effectively, rather than manually tuning.

5. REFERENCES

[1] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman, "From few to many: illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Analysis and Machine intelligence*, vol. 23, no. 6, pp. 643–659, 2001.

[2] J. Wright, A. Y. Yang, A. Ganesh, S. Ganesh, and S. S. Sastry, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, vol. 31, no. 2, pp. 1-18, 2009.

[3] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, "Robust face recognition for uncontrolled pose and illumination changes," *IEEE Transactions on Systems, Man and Cybernetics: Systems*, vol. 43, no. 1, pp. 149–163, 2013.

[4] A.A Shafie, F. Hafiz, and Y.M. Mustafah, "Face recognition using illumination-invariant local patches," *International Conference on Intelligent and Advanced Systems*, pp. 1–6, 2014.

[5] H. Wang, S. Li, Y. Wang, and J. Zhang, "Self quotient image for face recognition," *International Conference on Image processing, 2004.*, vol. pp. 1397-1400, 2004.

[6] T. Zhang, Y.Y. Tang, B. Fang, Z. Shang, and X. Liu, "Face recognition under varying illumination using gradientfaces," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2699–2606, 2009.

[7] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.

[8] C. Ding, J. Choi, D. Tao, and L.S. Davis, "Multi-directional multi-level dual-cross patterns for robust face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 3, pp. 518–531, 2016.

[9] T. Le, "On approaching heuristic weight mask to enhance lbp-based profile face recognition system," *Indian Journal of Science and Technology*, vol. 9, no. 17, pp. 1–7, 2016.

[10] A. Benzaoui, A. Boukrouche, H. Doghmane, and H. Bourouba, "Face recognition using 1dlbp, dwt and svm," *3rd International Conference on Control, Engineering & Information Technology*, pp. 1–6, 2015.

[11] P.P. Das, T. Jabid, and S.M.S Mahamud, "Single image face recognition based on gabor, sobel and local ternary pattern," *International Journal of Computer Applications*, vol. 132, no. 16, pp. 15–19, 2015.

[12] K-C. Fan and T-H. Hung, "A novel local pattern descriptor - local vector pattern in high-order derivative space for face recognition," *IEEE Transactions on Image Processing*, vol. 23, no. 7, pp. 2877–2891, 2014.

[13] B. Zhang, Y. Gao, S. Zhao, and J. Liu, "Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor," *Trans. Img. Proc.*, vol. 19, no. 2, pp. 533–544, 2010.

[14] L. Lenc and P. Král, "Automatically detected feature positions for lbp based face recognition," *Proceedings of the International Conference on Artificial Intelligence Applications and Innovations*, pp. 246–255, 2014.

[15] O. Nikisins, "Weighted multi-scale local binary pattern histograms for face recognition," *International Conference on Applied Mathematics and Computational Methods*, pp. 76–81, 2013.

[16] Q.-C. Tao, Z.-M. Liu, G. Bebis, and M. Hussain, "Face recognition using a novel image representation scheme and multi-scale local features," *International Journal of Biometrics*, vol. 7, no. 3, pp. 191–212, 2015.

[17] Z. Guo, L. Zhang, D. Zhang, and X. Mou, "Hierarchical multiscale lbp for face and palmprint recognition," *Proceedings of 2010 IEEE 17th International Conference on Image Processing*, pp. 4521–4524, 2010.

[18] N.-S. Vu, H.M. Dee, and A. Caplier, "Face recognition using the POEM descriptor," *Pattern Recognition*, vol. 45, pp. 2478–2488, 2012.

[19] M.X. Nguyen, M.L. Quang, V. Pham, T. Tran, and B.C. Le, "Multi-scale sparse representation for robust face recognition," *Third International Conference on Knowledge and Systems Engineering (KSE)*, pp. 195–199, 2011.

[20] P. Zhu, L. Zhang, Q. Hu, and S. Shiu, "Multi-scale patch based collaborative representation for face recognition with margin distribution optimization," *Computer Vision - ECCV 2012.*, vol. 7572, pp. 822-835, 2012.

[21] G. Gao, J. Yang, X. Jing, P. Huang, J. Hua, and D. Yue, "Robust face recognition via multi-scale patch-based matrix regression," *PLoS One*, vol. 11, no. 8, pp. 1–19, 2016.

[22] Y. Xu, X. Feng, X. Li, J. Yang, J. You, H. liu, and S. Teng, "Data uncertainty in face recognition," *IEEE Transactions on Cybernetics*, vol. 44, no. 10, pp. 1950–1961, 2014.

[23] Z.-Q. Zhao, Y.-M. Cheung, H. Hu, and X. Wu, "Expanding dictionary for robust face recognition: pixel is not necessary while sparsity is," *IET Computer Vision*, vol. 9, no. 5, pp. 648–654, 2015.

[24] D.-Y. Huang, C.-N. Huang, and W.-C. Hu, "Adaptable dense representation with collaborative representation classification for face recognition using low-rank matrices recovery," *International Conference on Advanced Information Networking and Applications Workshops*, pp. 902–907, 2016.

[25] S. Shan, W. Gao, B. Cao, and D. Zhao, "Lightness and retinex theory," *Journal of Optical Society of America*, vol. 61, no. 1, pp. 1–11, 1971.

[26] N. McLaughlin, J. Ming, and D. Crookes, "Illumination invariant facial recognition using a piecewise-constant lighting model," *2012 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1831–1834, 2012.

[27] C. Anderson, "LFWcrop face dataset," http://conradsanderson.id.au/lfwcrop/, [Online; accessed 9-February-2016].

[28] L. Lenc and P. Král, "Unconstrained Facial Images: Database for face recognition under real-world conditions," in *14th Mexican International Conference on Artificial Intelligence*, 2015, pp. 25–31.

[29] Brian Becker and Enrique Ortiz, "Evaluation of face recognition techniques for application to facebook," in *8th IEEE International Conference on Automatic Face & Gesture Recognition*, 2009.

[30] L. Wolf, T. Hassner, and Y. Taigman, "Effective unconstrained face recognition by combining multiple descriptors and learned background statistics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 10, pp. 1978–1990, 2010.

[31] D. Rim, M.K. Hasan, F. Puech, and C.J. Pal, "Learning from weakly labelled faces and video in the wild," *Pattern Recognition*, vol. 48, no. 3, pp. 759–771, 2015.